

# Thermoelectric performance of strongly-correlated quantum impurity models

Edward Taylor and Dvira Segal

*Chemical Physics Theory Group, Department of Chemistry,  
University of Toronto, 80 Saint George St. Toronto, Ontario, Canada M5S 3H6*

(Dated: September 23, 2015)

We derive asymptotically exact expressions for the thermopower and figure of merit of a quantum impurity connecting two noninteracting leads in the linear response regime where the chemical potential and temperature differences between the leads are small. Based on sum rules for the single-particle impurity spectral function, these expressions become exact at high temperatures as well as in the very strongly correlated regime, where the impurity Coulomb repulsion is much larger than the temperature. Although modest interactions impede thermoelectric performance, a very large Coulomb scale restores the optimal transport properties of noninteracting electrons, albeit renormalized to account for the absence of double occupancy in the impurity. As with noninteracting electrons, the electronic contribution to the figure of merit is limited only by the spectral broadening that arises from the coupling between the impurity and the leads.

PACS numbers: 73.50.Lw, 73.23.Hk, 85.65.+h

## I. INTRODUCTION

Nearly twenty years ago, Mahan and Sofo elucidated the properties that an ideal thermoelectric material—one that can efficiently transform a temperature differential into a voltage—should have<sup>1</sup>. They found that optimal thermoelectric conversion efficiency in the linear response regime, as determined by a large figure of merit, is realized in systems with a transmission function  $\mathcal{T}(E) \propto \delta(E - \epsilon_0)$  exhibiting a Dirac-delta function dependence on energy. In such a system, the electron contribution  $\kappa_{el}$  to the thermal conductance at zero charge current vanishes and the figure of merit

$$ZT = \frac{GS^2T}{\kappa_{el} + \kappa_{ph}} \quad (1)$$

is only limited by the smallness of  $\kappa_{ph}$ . Here  $G$  is the electrical conductance,  $S$  is the Seebeck coefficient (thermopower), and  $\kappa_{el}$  and  $\kappa_{ph}$  are the electrical and phonon contributions to the thermal conductance.

Although realistic systems do not exhibit such transmission, considerable experimental and theoretical work has concentrated on exploring thermoelectric properties of nearly single-level “impurity” systems such as quantum dots<sup>2–5</sup> or molecular junctions<sup>6–13</sup>. Such systems can in principle exhibit a transmission that is strongly peaked about this level and hence, promise large figures of merit. At the same time, these are *small* systems and their coupling to the leads has a significant effect on the transmission function  $\mathcal{T}(E)$ , yielding an extrinsic broadening  $\Gamma$ . Nonetheless, the phonon contribution to the thermal conductance is expected to be small in these systems as compared to that in bulk ones, and there is still great interest in exploring their thermoelectric properties. (Recent studies have stressed the importance of the phonon contribution  $\kappa_{ph}$  in molecules, however<sup>14,15</sup>.)

Beyond numerical studies of the thermoelectric performance of impurity models based on numerical renormalization group<sup>3,9,16</sup> and the nonequilibrium Green’s func-

tion approach<sup>17,18</sup>, simple analytic results for thermoelectric coefficients have been found in the “atomic limit”, in which the broadening  $\Gamma \rightarrow 0$  is taken to be zero, using e.g., the sequential-tunnelling approximation for transport<sup>19–21</sup>. In this limit, taking the Coulomb repulsion  $U$  to be zero, one recovers the idealized, un-broadened single-level limit of Mahan and Sofo, and correspondingly, an infinite electronic contribution to the figure of merit (i.e., after setting  $\kappa_{ph} = 0$ ). Turning on interactions, a second satellite level arises due to the interaction energy shift  $U$  when two electrons are present in the impurity. In the atomic limit  $\Gamma \rightarrow 0$  with strong interactions,  $U \gg T$ , Murphy, Mukerjee, and Moore<sup>21</sup> showed that the second level becomes unoccupied and hence, irrelevant for transport, meaning that the system again effectively reduces to a single-level one with a diverging electronic figure of merit. For intermediate coupling strength,  $U \sim T$ , both levels are active in transport and the figure of merit is not large in general.

Analogous studies of thermoelectricity in the atomic limit of bulk systems have been undertaken in a pair of well-known papers by Beni<sup>22</sup> and Chaikin and Beni<sup>23</sup> (see also Ref. 24). As emphasized by Beni<sup>22</sup>, however, it is nontrivial to perturb away from the atomic limit of transport since transport formally vanishes when  $\Gamma = 0$  (for them, the limit where the hopping matrix element between lattice sites vanishes). For a single-level spin-full impurity, the available Hilbert space for transport formally evolves from having three elements, corresponding to the three possible occupancies, to an infinite number, a large fraction of which must be used in order to have a conserving theory of transport<sup>25</sup>. Even when interactions are modest, it is challenging to develop a reliable perturbation expansion of transport quantities in powers of the coupling matrix elements between the impurity and leads (effectively,  $\Gamma$ ).

In this paper, we use sum rules for the impurity electron spectral function to derive expressions for the thermoelectric coefficients that are asymptotically exact

in two regimes for which the broadening is small, but nonzero: First, at high temperatures, greater than  $U, \Gamma$  and the bandwidth  $D$  that characterizes the leads. Second, in the very strongly-correlated, “narrow-level” limit where  $U \gg T \gg \Gamma$  and  $U > D$ . Our results in this latter regime reproduce the divergent figure of merit in the atomic limit when  $U \rightarrow \infty^{21}$ . At the same time, this limit is highly singular and we find

$$\lim_{\Gamma \rightarrow 0} ZT(U \gg T) \sim \frac{(\epsilon_0 - \mu)^2}{\Gamma D}, \quad (2)$$

where  $\epsilon_0 - \mu$  is the difference between the impurity energy and Fermi levels. This result emphasizes the difficulty of studying transport in this limit and also the crucial role played by the broadening in the thermoelectric performance of impurity models with strong interactions.

We start in Sec. II by introducing the linear-response formulae for thermoelectric transport coefficients in terms of the single-particle impurity spectral function. In Sec. III, we introduce two sum rules for this spectral function—one that integrates all spectral weight and another which removes the irrelevant weight in the upper Hubbard peak—and show how these can be used to derive results for transport that are asymptotically exact in the regimes elucidated above. We then calculate explicit forms for these sum rules for the Anderson impurity model in Sec. IV, and use these results to discuss thermoelectric performance in Secs. V and VI. Finally, in Sec. VII, we summarize our main results and conclude with a discussion of the implications of our results for quantum dot and molecular junction systems.

## II. TRANSPORT COEFFICIENTS

Our starting point is the (generalized) Landauer-like expressions for the charge<sup>26</sup>

$$J = \frac{e}{2\pi} \int_{-\infty}^{\infty} dE \mathcal{T}(E, T_L, T_R) [f_L(E) - f_R(E)] \quad (3)$$

and heat

$$J_Q = \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \mathcal{T}(E, T_L, T_R) (E - \mu_L) [f_L(E) - f_R(E)] \quad (4)$$

currents through an impurity connecting two leads.  $f_\nu(E) \equiv \{\exp[\beta_\nu(E - \mu_\nu)] + 1\}^{-1}$  is the Fermi function at the left ( $\nu = L$ ) and right ( $\nu = R$ ) leads, with temperature  $T_\nu \equiv \beta_\nu^{-1}$  and chemical potential  $\mu_\nu$ . Unless specified otherwise, throughout this paper we set  $\hbar = k_B = 1$ . When the coupling between the impurity and the left and right leads are proportional, Meir and Wingreen showed that (3) is an exact expression with the identification of the generalized transmission function with the spectral function  $A(E)$  of the impurity electrons<sup>27</sup>:

$$\mathcal{T}(E, T_L, T_R) = (\pi\gamma/2)\Gamma(E)A(E - \mu), \quad (5)$$

where

$$\Gamma \equiv \Gamma_L + \Gamma_R \quad (6)$$

is the sum of the broadenings [assumed in (5) to be purely real] at the left and right leads,

$$\gamma \equiv 4\Gamma_L\Gamma_R/\Gamma^2 \quad (7)$$

is an asymmetry parameter that deviates from unity when  $\Gamma_L \neq \Gamma_R$ , and

$$A(E) \equiv -\frac{1}{\pi} \sum_{\sigma} \text{Im}G_{\sigma}^{\text{ret}}(E) \quad (8)$$

is the spectral function for both spin species  $\sigma$  of impurity electrons with retarded Green’s function  $G^{\text{ret}}(E)$ . It trivially follows from the analysis in Ref. 27 that (4) is also exact with the identification in (5)<sup>9,28</sup>.

The precise form of the spectral function  $A(E)$  is specified by the details of the coupling between the impurity and the leads. To be specific, suppose

$$\hat{V} = \sum_{\mathbf{k}\nu\sigma} [V_{\mathbf{k}\nu} \hat{c}_{\mathbf{k}\nu\sigma}^{\dagger} \hat{d}_{\sigma} + \text{h.c.}] \quad (9)$$

couples the single-level impurity electrons characterized by  $\hat{d}_{\sigma}$  to the non-interacting  $\nu = L, R$  leads, and

$$\hat{H}_l = \sum_{\mathbf{k}\nu\sigma} (\epsilon_{\mathbf{k}\nu} - \mu_{\nu}) \hat{n}_{\mathbf{k}\nu\sigma} \quad (10)$$

describes noninteracting electrons with momentum distribution  $\hat{n}_{\mathbf{k}\nu\sigma} \equiv \hat{c}_{\mathbf{k}\nu\sigma}^{\dagger} \hat{c}_{\mathbf{k}\nu\sigma}$  in these leads. For this model,  $A(E)$  is given by

$$A(E) = \sum_{\sigma} \sum_{a,b} (P_a + P_b) \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle \langle a | \hat{d}_{\sigma} | b \rangle \delta(E - E_b + E_a), \quad (11)$$

and the broadening is

$$\Gamma_{\nu}(E) = 2\pi \sum_{\mathbf{k}} \delta(E - \epsilon_{\mathbf{k}\nu}) |V_{\mathbf{k}\nu}|^2. \quad (12)$$

The states  $|a\rangle, |b\rangle$  in the spectral representation (11) are the exact eigenstates of the many-body grand-canonical hamiltonian,  $\hat{H}|a\rangle = E_a|a\rangle$ . Within linear response  $\mu_L \simeq \mu_R$  and  $T_L \simeq T_R$ ,  $P_a \equiv \exp(-\beta E_a)/\mathcal{Z}$ , with  $\mathcal{Z}$  the grand canonical partition function.  $\hat{H}$  includes (9) and (10) as well as the as-yet-unspecified impurity hamiltonian. We will restrict ourselves in what follows to the linear response regime.

Equations (3), (4), and (5) lead to the following expressions for the transport coefficients<sup>1,26</sup> in the linear response regime for an energy-independent broadening  $\Gamma_{\nu}(E) = \Gamma_{\nu}$  (momentarily restoring  $\hbar$  and  $k_B$ ):

$$G = \frac{e^2 \gamma \Gamma}{16\hbar k_B T} M_0, \quad (13)$$

$$\kappa_{el} = \frac{\gamma \Gamma}{16\hbar k_B T^2} \left( M_2 - \frac{M_1^2}{M_0} \right), \quad (14)$$

$$S = \frac{1}{eT} \frac{M_1}{M_0}, \quad (15)$$

and

$$ZT = \frac{M_1^2}{M_0(M_2 - M_1^2/M_0)}, \quad (16)$$

defined in terms of the integral expression

$$M_n \equiv \int_{-\infty}^{\infty} dE E^n \operatorname{sech}^2(\beta E/2) A(E). \quad (17)$$

In arriving at these expressions, we have shifted  $E \rightarrow E + \mu$  in the linear-response limit  $\mu_L = \mu_R = \mu$  of the expressions (3) and (4) for the currents, meaning that the spectral function  $A(E)$  is the one defined in (11). An energy-independent broadening is appropriate for leads that exhibit good metallic behaviour with a broad bandwidth, generally much larger than  $\Gamma$ . Had we not made this assumption, the broadening  $\Gamma(E)$  would have entered the integral expression (17). (As we explain below, this situation can also be dealt with using the methods developed in this paper, although the resulting calculations would be more complicated.) The above expressions will form the basis of our analysis in the remainder of this paper.

Before closing this section, we briefly comment on the major result of Mahan and Sofo who, as noted earlier, showed that the electronic contribution to the figure of merit would diverge in a system described by a single, un-broadened energy level. In the language of this section, this corresponds to having a Dirac-delta function spectral function

$$A(E) \propto \delta(E - \epsilon_0). \quad (18)$$

Using this in the above, the thermopower is (restoring  $k_B$ )  $S = (k_B/e)(\epsilon_0/T)$ . Crucially, using (18) in (17), one finds

$$M_2 = \frac{M_1^2}{M_0}, \quad (19)$$

and the electronic contribution to the thermal conductance vanishes identically at *all temperatures*, with the result that the figure of merit is divergent at all temperatures.

Moving away from the limit (18) of having a single un-broadened energy level, a small thermal conductance can more generally be understood as expressing a small variance in the single-particle energy. Defining

$$\langle\langle \dots \rangle\rangle_F \equiv \frac{\int dE (\dots) A(E) F(E)}{\int dE A(E) F(E)}, \quad (20)$$

we see that

$$M_2 - \frac{M_1^2}{M_0} = M_0 [\langle\langle E^2 \rangle\rangle_F - \langle\langle E \rangle\rangle_F^2], \quad (21)$$

with  $F = \operatorname{sech}^2(\beta E/2)$ . Hence, the ratio

$$\frac{\kappa_{el}}{GT} = \frac{1}{e^2 T^2} (\langle\langle E^2 \rangle\rangle_F - \langle\langle E \rangle\rangle_F^2)_{F=\operatorname{sech}^2(\beta E/2)}, \quad (22)$$

which defines the Lorenz number in bulk systems, provides a direct measure of the variance in the single-particle energies clustered within  $T$  of the Fermi level (because of the form of  $F$ ). It is thus clear why a system with a single un-broadened energy level would be optimal, having zero variance in its energy levels; see also Eq. (18) in Mahan and Sofo<sup>1</sup>. This result also presages our main conclusion: Even though an extremely large  $U$  can eliminate the energy-level variance due to interactions, unless the broadening  $\Gamma$  due to coupling with the leads can be made smaller than the temperature (not currently the situation in experiments<sup>6,29</sup>), the energy variance will be greater than  $T^2$ , and one should only expect modest thermoelectric performance.

### III. SUM RULE EXPRESSIONS FOR TRANSPORT COEFFICIENTS

The evaluation of the spectral function  $A(E)$  that enters the expressions for the thermoelectric coefficients is a challenging many-body problem. In this paper, we propose an alternative, evaluating instead

$$M_n^T \equiv \int_{-\infty}^{\infty} dE E^n A(E) \quad (23)$$

and

$$\tilde{M}_n(\Omega_c) \equiv \int_{-\infty}^{\Omega_c} dE E^n A(E). \quad (24)$$

The first of these is a well-known *sum rule*, and can straightforwardly be evaluated in terms of commutators involving electron creation and annihilation operators and the hamiltonian (see e.g., Refs. 30,31). Integrating the product of  $E^n$  and (11) leads to (see Appendix A)

$$M_n^T = \sum_{\sigma} \langle \{ \hat{d}_{\sigma}^{\dagger}, [\hat{d}_{\sigma}, \hat{H}]_n \} \rangle. \quad (25)$$

Here  $[\hat{d}_{\sigma}, \hat{H}]_n$  is a nested commutator with  $[\hat{d}_{\sigma}, \hat{H}]_0 = \hat{d}_{\sigma}$  the zeroth-order commutator,  $[\hat{d}_{\sigma}, \hat{H}]_1 = [\hat{d}_{\sigma}, \hat{H}]$  the first-order commutator,  $[\hat{d}_{\sigma}, \hat{H}]_2 = [[\hat{d}_{\sigma}, \hat{H}], \hat{H}]$ , and so on.

The second expression, (24), involving an energy cutoff  $\Omega_c$ , is evaluated by projecting out the states with energy above  $\Omega_c$ . Taking  $\Gamma \ll \Omega_c \ll U$ , this amounts to deriving an effective operator  $d_{\sigma, \epsilon_0}$  for the lower Hubbard peak. The resulting sum rule for the spectral weight contained in the lower Hubbard peak of a single-level impurity model is

$$\tilde{M}_n = \sum_{\sigma} \langle \{ d_{\sigma, \epsilon_0}^{\dagger}, [d_{\sigma, \epsilon_0}, \hat{H}]_n \} \rangle. \quad (26)$$

The precise choice of operator  $d_{\sigma, \epsilon_0}$  depends on the model under consideration; its derivation for the Anderson impurity model coupled to two leads will be given in the next section. We could equally derive a sum rule for the spectral weight contained in the upper Hubbard peak,

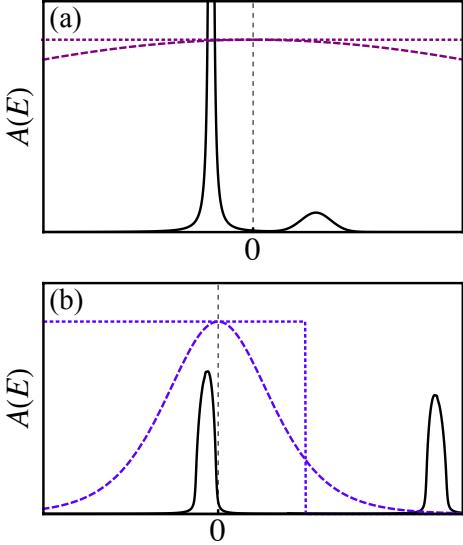


FIG. 1: Schematic plot of the spectral function  $A(E)$  at high temperatures  $T \gg U, \Gamma$  (a) and for strong correlations  $\Gamma \lesssim T \ll U$  (b) relative to the lead Fermi level (denoted here by  $E = 0$ ). Also shown by the dashed lines is the thermal weight function  $\text{sech}^2(\beta E/2)$  that enters the moments  $M_n$ ; the dotted lines show the weight functions—unity and the step function  $\Theta(\Omega_c - E)$ —involved with the sum rules.

as would be relevant for the situation when the chemical potential is close to this peak. In what follows, we will assume that the Fermi level is close to the lower Hubbard peak, however.

The sum rule (23) provides a rigorous upper bound on  $M_n$  when  $n$  is even:

$$M_n \leq M_n^T \quad n = 0, 2, \dots \quad \forall T, \quad (27)$$

since  $A(E) \geq 0 \quad \forall E$  and  $\text{sech}^2(x) \leq 1 \quad \forall x$ . This immediately allows us to write down upper bounds on the charge conductance

$$G \leq \frac{e^2 \Gamma}{8 \hbar T} M_0^T \quad (28)$$

and, making use of the fact that  $M_1^2/M_0 \geq 0$ <sup>32</sup>, thermal conductance  $\kappa_{el}$  at zero charge current:

$$\kappa_{el} \leq \frac{\Gamma}{2 \hbar T^2} M_2^T. \quad (29)$$

Related bounds for the thermal conductance due to bosonic excitations (e.g., phonons) were derived in Ref. 33.

Because they involve a ratio of moments, (27) does not obviously translate to bounds for the Seebeck coefficient  $S$  and figure of merit  $ZT$ . The usefulness of the sum rules is that they provide asymptotically exact approximations to the transport coefficients at high-temperatures, including the asymptotic limit  $T \gg U, \Gamma, D$ , and also in the strongly-correlated regime  $\Gamma \lesssim T \ll U$ . In Fig. 1, we

show two schematic plots of the impurity spectral function along with the weight function  $\text{sech}^2(\beta E/2)$  that enters the exact transport moments  $M_n$ . In Fig. 1(a), the temperature is much larger than the Coulomb repulsion  $U$  as well as  $\Gamma$ ; the majority of the spectral weight arises in a Lorentzian centred around the Fermi level of width  $\sim \Gamma$ . A small peak—the weight of which vanishes with decreasing  $U/T$ —arises at  $U$  corresponding to the energy of a doubly occupied impurity. In this regime, the  $\text{sech}^2(\beta E/2)$  function is essentially unity wherever there is nonzero spectral weight and one can replace  $M_n$  by the high-temperature asymptotes  $M_n^T$ , with the result that the thermoelectric transport coefficients calculated using these sum rules become asymptotically exact:

$$\lim_{T \gg U, \Gamma, D} \{G, \kappa_{el}, S, ZT\} [M_n] = \{G, \kappa_{el}, S, ZT\} [M_n^T]. \quad (30)$$

We formally require that  $T \ll D$  be much larger than the lead bandwidth since the spectral function scales as  $A(E) \propto E^{-2}$  for large  $E$ , meaning that  $M_1^T$  and  $M_2^T$  cannot well-approximate  $M_1$  and  $M_2$  unless  $D \lesssim T$ .

Fig. 1(b) shows the spectral function in the strongly-correlated regime  $\Gamma \ll T \ll U$ . Here two Hubbard peaks develop, the lower one centred at some single-particle energy scale,  $\epsilon_0$ , and the upper one centred at  $\epsilon_0 + U$ . When the Fermi level (denoted by the zero in these plots) is within  $\sim T$  of e.g. the lower Hubbard peak, the  $\text{sech}^2(\beta E/2)$  function excludes the upper Hubbard peak. Because there is negligible spectral weight outside the two peaks when  $U \gg \Gamma$ <sup>34</sup>, in this regime  $M_n$  can be replaced by  $\tilde{M}_n(\Omega_c)$ , where  $\Gamma \ll \Omega_c \ll U$  and the transport coefficients again admit asymptotically exact sum rule values:

$$\lim_{\Gamma \ll T \ll U} \{G, \kappa_{el}, S, ZT\} [M_n] = \{G, \kappa_{el}, S, ZT\} [\tilde{M}_n(\Gamma \ll \Omega_c \ll U)]. \quad (31)$$

Even though (31) is asymptotically exact in the limit  $\Gamma \ll T$ , it is evident that these results will remain qualitatively valid for  $\Gamma \lesssim T$ .

As noted earlier, we could also consider the thermoelectric performance when the Fermi level is close to the upper Hubbard peak by considering the sum rule for this peak. Apart from effects that break particle-hole symmetry (enhanced by molecular vibrational modes<sup>18</sup>), the thermoelectric performance will be the same in both cases.

In Figure 2, we plot the regimes in the parameter space spanned by  $U/T, \Gamma/T$  where our results hold. Also shown is the regime studied by Murphy, Mukerjee, and Moore<sup>21</sup>, as well as that studied using NRG by Costi and Zlatić<sup>3</sup>, although NRG can study the entire region shown in this figure. Even though it might seem that we are dealing with a highly restricted region of parameter space—and one that is moreover far from most current experiments involving molecular junctions for which  $\Gamma \sim \mathcal{O}(10T)$ <sup>6,29</sup>—we reemphasize that it is *only when  $\Gamma \ll T$  that a substantial figure of merit can be realized*. We also note that

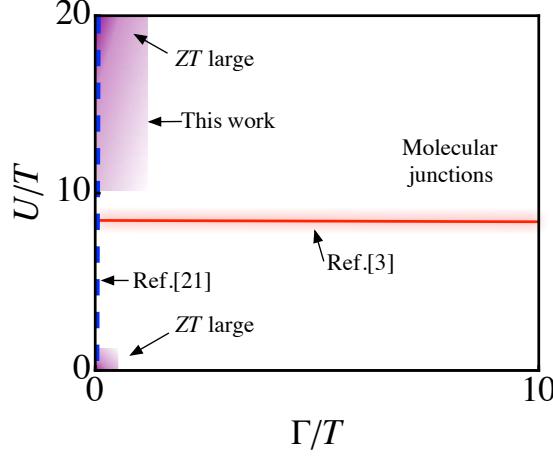


FIG. 2: Plot of the regimes where our results hold ( $U \gg T \gtrsim \Gamma$  and  $T \gg U, \Gamma, D$ ), in comparison to the regimes studied in the sequential tunnelling approximation of Ref. 21 (all  $U, \Gamma/T \rightarrow 0$ ) and the numerical renormalization group calculations of Ref. 3 ( $U/\Gamma = 8$ , all  $\Gamma$ ). Note that the latter calculation could have been extended to any region shown above. Also indicated are the regimes ( $\Gamma, U \ll T$  and  $\Gamma \ll T, U \gg T$ ) where the electronic contribution to the figure of merit becomes large as well as the parameter regime of typical molecular junction experiments. Although the value of  $U/T$  is unknown,  $\Gamma$  is typically  $\gtrsim 10T$ <sup>6,29</sup>.

this regime is well away from the Kondo regime  $T \lesssim T_K$ , since  $T_K \ll \Gamma$ <sup>3,35</sup>

In the remainder of this paper, we investigate the implications of (30) and (31) for an archetypal model of strongly correlated “impurity” electrons, the Anderson impurity model, generalized to include the coupling between an impurity and two leads. These sum-rule results have been derived assuming a constant broadening function, as is almost always taken to be the case. One can also consider energy-dependent broadenings: for simple power-law forms such as  $\Gamma(E) = \Gamma E^l$  with  $l$  and integer, the resulting transport coefficients simply involve higher-moment sum rules ( $n \rightarrow n + l$ ). For more complicated dependencies, one must resort to the operator-product expansion technique<sup>36</sup> of Wilson, Kadanoff, and Polyakov.

#### IV. SUM RULES FOR THE ANDERSON IMPURITY MODEL

The single-level Anderson impurity model<sup>35</sup> (AIM) for an impurity coupled to two leads is

$$\hat{H} = \hat{H}_d + \hat{H}_l + \hat{V}. \quad (32)$$

Here

$$\hat{H}_d = (\epsilon_0 - \mu) \sum_{\sigma} \hat{d}_{\sigma}^{\dagger} \hat{d}_{\sigma} + U \hat{n}_{\uparrow} \hat{n}_{\downarrow} \quad (33)$$

is the impurity hamiltonian and  $H_l$  and  $\hat{V}$  are given by (10) and (9), the former with  $\mu_{\nu} = \mu$ , as appropriate in the linear response regime.

#### A. Sum rules for the entire spectrum

The results of a straightforward calculation using (25) and (32) are

$$M_0^T = 2, \quad (34)$$

$$M_1^T = 2(\epsilon_0 - \mu + Un_d/2), \quad (35)$$

and

$$M_2^T = 2(\epsilon_0 - \mu)^2 + 2(\epsilon_0 - \mu)Un_d + U^2n_d + 2 \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2. \quad (36)$$

Here  $n_d \equiv \sum_{\sigma} n_{\sigma}$  is the total occupancy of the impurity. The factor of two in (34) (and elsewhere in the above) counts the number of single-particle states: two for the single-level AIM with two spins. Taking the  $V_{\mathbf{k}\nu}$  to be independent of momentum,

$$\begin{aligned} \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 &= \sum_{\nu} |V_{\nu}|^2 N_0 \int_{-D}^D dE \\ &= 2DN_0 \sum_{\nu} |V_{\nu}|^2. \end{aligned} \quad (37)$$

Here we have taken the lead DOS  $N_{\sigma}(E) = N_0$  per spin to be constant (and independent of spin) with  $D$  the half-bandwidth. Making the same assumptions to evaluate the broadening (32), it becomes

$$\Gamma_{\nu} = 2\pi N_0 |V_{\nu}|^2, \quad (38)$$

and thus

$$\sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 = D\Gamma/\pi. \quad (39)$$

As noted in (21), the combination  $M_2^T - (M_1^T)^2/M_0^T = M_0^T [\langle\langle E^2 \rangle\rangle_{F=1} - \langle\langle E \rangle\rangle_{F=1}^2]$  provides a measure of the variance in the single-particle energies, in this case over the entirety of the spectral weight since there is no thermal weighting function  $\text{sech}^2(\beta E/2)$ . In order to have a small electronic contribution to the thermal conductance and hence, a substantial figure of merit, the energy variance in the system needs to be small. Equations (34)-(36) give

$$M_2^T - \frac{(M_1^T)^2}{M_0^T} = \frac{U^2}{2} n_d (2 - n_d) + 2 \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2. \quad (40)$$

The first term on the right-hand side gives the variance in the interaction energy per particle (proportional to  $U^2$  times the compressibility), while the second term describes the variance in the kinetic energy per particle. One can understand this result in terms of Fig. 1: with two peaks separated by  $U$ , the energy variance is clearly  $\sim U^2$ . The variance in a *single band* (either one) however, is  $\sim \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 \sim \Gamma D$ .

At temperatures greater than both  $\Gamma$  and  $U$ , the variance is of course large, and (36) indicates a correspondingly large  $\kappa_{el}$ . On the other hand, for lower temperatures,  $T \ll U$ , assuming the Fermi level is close to the

lower Hubbard peak, the  $\text{sech}^2(\beta E)$  factor in  $M_n$  is sensitive only to the spectral weight in that peak. Correspondingly,  $\kappa_{el}$  is only sensitive to the variance  $\sim \Gamma D$  in this peak. This notion can be made precise by projecting out the spectral weight contained in the upper Hubbard peak from the sum rules, as in (24). We turn to this now.

## B. Projected sum rules for the lower Hubbard peak

Relevant to the case shown in Fig. 1(b), we now calculate the sum rule for the lower Hubbard peak (LHP), obtained by projecting out the doubly-occupied states that give rise to the upper Hubbard peak (UHP). Our approach here follows closely that used in studies of the bulk electron Hubbard model<sup>31,37,38</sup> and also the AIM itself by Schrieffer and Wolff<sup>39</sup>. To remove the spectral weight associated with the UHP, we need a spectral decomposition for the impurity electron operator of the form

$$d_\sigma = d_{\sigma, \epsilon_0} + d_{\sigma, \epsilon_0 + U}. \quad (41)$$

Here  $d_{\sigma, E}$  describes an annihilation process that lowers the impurity energy by  $\sim E$  modulo  $\Gamma$ , i.e.,  $\langle b | d_{\sigma, E} | a \rangle = 0$  unless  $E_a - E_b = E + \mathcal{O}(\Gamma)$ . Once a suitable expression is found for  $d_{\sigma, \epsilon_0}$ , one can proceed to calculate the projected sum rules given by (26).

When  $\Gamma = 0$ , the decomposition is particularly simple since the occupancies  $\hat{n}_\sigma$  are good quantum numbers, and, defining  $\bar{\sigma} = -\sigma$ ,

$$d_{\sigma, \epsilon_0} = (1 - \hat{n}_{\bar{\sigma}}) \hat{d}_\sigma \quad (42)$$

and

$$d_{\sigma, \epsilon_0 + U} = \hat{n}_{\bar{\sigma}} \hat{d}_\sigma. \quad (43)$$

When  $\Gamma \neq 0$ , we seek a *renormalized* operator  $\tilde{d}_\sigma$  such that  $\tilde{n}_\sigma$  is a good quantum number of the bare hamiltonian  $\hat{H}$ , and  $[\tilde{n}_\sigma, \hat{H}] = 0$ . This allows us to undertake a similar spectral decomposition as (41) and identify the renormalized operator  $d_{\sigma, \epsilon_0}$  for the LHP needed for the sum rules.

Since  $\tilde{n}_\sigma$  commutes with  $\hat{H}_0 \equiv \hat{H}_0(\hat{d}_\sigma \rightarrow \tilde{d}_\sigma, \hat{c}_{\mathbf{k}\nu\sigma} \rightarrow \tilde{c}_{\mathbf{k}\nu\sigma})$ , where

$$\hat{H}_0 \equiv \hat{H}_d - \mu \sum_{\mathbf{k}\nu} \hat{n}_{\mathbf{k}\nu\sigma} \quad (44)$$

is the total hamiltonian less all kinetic energies,  $[\tilde{n}_\sigma, \hat{H}] = 0$  implies  $[\tilde{H}_0, \hat{H}] = 0$ . Relating the renormalized operators to the original operators by a unitary transformation (not to be confused with the thermopower  $S$ , but in keeping with widely-used notation)<sup>31,37–40</sup>  $\hat{d}_\sigma = e^S \tilde{d}_\sigma e^{-S}$  and  $\hat{c}_{\mathbf{k}\nu\sigma} = e^S \tilde{c}_{\mathbf{k}\nu\sigma} e^{-S}$ , this last condition amounts to finding a transformation operator  $S$  that satisfies

$$[\hat{H}, \tilde{H}_0] = 0 \Rightarrow [e^S \tilde{H} e^{-S}, \tilde{H}_0] = 0. \quad (45)$$

Here, and in the remaining, we use a tilde to indicate a transformed operator.

The fact that  $\tilde{n}_\sigma$  is a good quantum number means that—analogous to (41)—we can decompose any operator  $\tilde{O} = \tilde{O}(\tilde{d}_\sigma, \tilde{c}_{\mathbf{k}\nu\sigma})$  in terms of the eigenstates  $\tilde{n}_d = \sum_\sigma \tilde{n}_\sigma = 0, 1, 2$ , or equivalently, the corresponding energy changes  $E = \pm\epsilon_0, \pm(\epsilon_0 + U)$  that they effect, modulo corrections  $\mathcal{O}(\Gamma)$ :

$$\tilde{O} = \tilde{O}_{\pm\epsilon_0} + \tilde{O}_{\pm(\epsilon_0 + U)}. \quad (46)$$

$\tilde{O}_{\pm E}$  operators are related by hermitian conjugacy:

$$\tilde{O}_{-E} = \tilde{O}_E^\dagger. \quad (47)$$

Equation (41) is a special example of this more general spectral decomposition. Because the eigenvalues  $\tilde{n}_d$  of  $\tilde{H}_0$  are good quantum numbers, so are the energy changes  $E$  defined above, with the result that

$$[\tilde{O}_E, \tilde{H}_0] = E \tilde{O}_E + \mathcal{O}(\Gamma), \quad E = \pm\epsilon_0, \pm(\epsilon_0 + U). \quad (48)$$

Equation (48) allows us to determine  $S$  and  $d_{\sigma, \epsilon_0}$  to a specified order in  $V/(\epsilon_0 + U)$ , where  $V$  is the characteristic size of the in-general momentum dependent coupling  $V_{\mathbf{k}\nu}$ . We expand  $\hat{H} = \exp(S) \tilde{H} \exp(-S)$  as  $\hat{H} = \tilde{H} + [S, \tilde{H}] + \dots$  and  $S = S^{(1)} + S^{(2)} + \dots$  in powers of  $V/(\epsilon_0 + U)$  and use this in (45). Requiring that  $S$  eliminates the  $\tilde{H}_{\pm(\epsilon_0 + U)}$  terms in the spectral decomposition (46) of  $\hat{H}$  (thereby removing the associated spectral weight) and using (48), (45) reduces to

$$[S^{(1)}, \tilde{H}_d] = -\tilde{T}_{\epsilon_0 + U} - \tilde{T}_{-\epsilon_0 - U}. \quad (49)$$

Here  $\tilde{T}_E$  are the terms in the spectral decomposition of the kinetic energy contribution  $\tilde{T} \equiv \tilde{H} - \tilde{H}_0$  to the hamiltonian; note that these are the only terms in the hamiltonian that can change the impurity energy by  $\sim \pm(\epsilon_0 + U)$ :  $\tilde{H}_{\pm(\epsilon_0 + U)} = \tilde{T}_{\pm(\epsilon_0 + U)}$ . Explicitly,

$$\tilde{T}_{\epsilon_0} = \sum_{\mathbf{k}\nu\sigma} \epsilon_{\mathbf{k}\nu} \tilde{n}_{\mathbf{k}\nu\sigma} + \sum_{\mathbf{k}\nu\sigma} [V_{\mathbf{k}\nu} \tilde{c}_{\mathbf{k}\nu\sigma}^\dagger \tilde{d}_\sigma (1 - \tilde{n}_{\bar{\sigma}}) + \text{H.c.}] \quad (50)$$

and

$$\tilde{T}_{\epsilon_0 + U} = \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* \tilde{n}_{\bar{\sigma}} \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma}. \quad (51)$$

Using (51) and (47) in (49), one finds

$$S^{(1)} = -\frac{1}{\epsilon_0 + U} [\tilde{T}_{-\epsilon_0 - U} - \tilde{T}_{\epsilon_0 + U}]. \quad (52)$$

This is essentially<sup>41</sup> just the Schrieffer–Wolff transformation used by Schrieffer and Wolff to derive the Kondo hamiltonian as the effective low-energy description of the AIM<sup>39</sup>. Apart from the presence of the additional single-particle energy scale  $\epsilon_0$  in this expression, it is also the same result as obtained for the bulk Hubbard model, with

$\tilde{V}$  replaced by inter-site tunnelling; see e.g., Eq. (7) in Ref. 40 and Eq. (10) in Ref. 37.

Having determined the form of  $S^{(1)}$ , we can now obtain an expression for the renormalized LHP operator appearing in (41). By expanding  $\hat{d}_\sigma = \exp(S)\tilde{d}_\sigma \exp(-S) = \tilde{d}_\sigma + [S^{(1)}, \tilde{d}_\sigma] + \dots$  and taking the spectral decomposition of both sides of the resulting expression, one finds perturbative expressions for the spectral decomposition  $d_{\sigma, \epsilon_0}$  and  $d_{\sigma, \epsilon_0+U}$  of the *original* operator; c.f. (41). Using (49) and equating terms of the same order in  $E$  gives

$$d_{\sigma, \epsilon_0} = (1 - \tilde{n}_\sigma)\tilde{d}_\sigma + \frac{1}{\epsilon_0 + U} [\tilde{T}_{\epsilon_0+U}, \tilde{d}_{\sigma, -\epsilon_0-U}] - \frac{1}{\epsilon_0 + U} [\tilde{T}_{-\epsilon_0-U}, \tilde{d}_{\sigma, -\epsilon_0-U}], \quad (53)$$

for the LHP operator. Explicit evaluation of this expression results in

$$d_{\sigma, \epsilon_0} = (1 - \tilde{n}_\sigma)\tilde{d}_\sigma - \frac{1}{\epsilon_0 + U} (\tilde{n}_\sigma + \tilde{n}_{\bar{\sigma}})\tilde{d}_\sigma \sum_{\mathbf{k}\nu} V_{\mathbf{k}\nu}^* \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\bar{\sigma}} - \frac{1}{\epsilon_0 + U} \sum_{\mathbf{k}\nu} V_{\mathbf{k}\nu}^* \tilde{n}_\sigma \tilde{c}_{\mathbf{k}\nu\sigma} + \mathcal{O}[V/(\epsilon_0 + U)]^2. \quad (54)$$

In the limit  $U \rightarrow \infty$ , (54) reduces to (42).

Using (54) to evaluate the first few sum rules (26), making use of the fact that there are no doubly-occupied states in the transformed basis,  $\tilde{d}_\sigma^\dagger \tilde{d}_\sigma = 0$ , one finds after some straightforward but laborious algebra

$$\tilde{M}_0 = 2 - n_d - \frac{1}{\epsilon_0 + U} \sum_{\mathbf{k}\nu\sigma} [V_{\mathbf{k}\nu}^* \langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle + \text{H.c.}] + \mathcal{O}\left(\frac{V}{\epsilon_0 + U}\right)^2, \quad (55)$$

$$\tilde{M}_1 = (\epsilon_0 - \mu)\tilde{M}_0 + \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* \langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle + \mathcal{O}\left(\frac{V^2}{\epsilon_0 + U}\right), \quad (56)$$

and

$$\begin{aligned} \tilde{M}_2 &= (\epsilon_0 - \mu)\tilde{M}_1 + (2 - n_d) \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 \\ &- \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* (\epsilon_{\mathbf{k}\nu} - \mu) \langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle \\ &+ \left\langle \left( \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \right) \left( \sum_{\mathbf{k}'\nu'\sigma'} V_{\mathbf{k}'\nu'}^* \tilde{c}_{\mathbf{k}'\nu'\sigma'}^\dagger \tilde{d}_{\sigma'} \right) \right\rangle + \mathcal{O}\left(\frac{V^3}{\epsilon_0 + U}\right). \end{aligned} \quad (57)$$

The analogue of (55) for the bulk Hubbard model has been derived in Refs. 31,37,38. Note that e.g.

$$\sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* \langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle = \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu} \langle \tilde{c}_{\mathbf{k}\nu\sigma}^\dagger \tilde{d}_\sigma \rangle \quad (58)$$

is purely real in equilibrium, as required to have no charge or heat current. Defining

$$\hat{\mathcal{V}} \equiv \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu} \hat{c}_{\mathbf{k}\nu\sigma}^\dagger \hat{d}_\sigma, \quad (59)$$

such that  $\hat{V} \equiv \hat{\mathcal{V}} + \hat{\mathcal{V}}^\dagger$  is the lead-impurity coupling (9), these give the following result for the variance in the LHP:

$$\begin{aligned} \tilde{M}_2 - \frac{\tilde{M}_1^2}{\tilde{M}_0} &= (2 - n_d) \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 - \sum_{\mathbf{k}\nu\sigma} V_{\mathbf{k}\nu}^* (\epsilon_{\mathbf{k}\nu} - \mu) \langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle \\ &+ \langle \tilde{\mathcal{V}}^2 \rangle - \frac{1}{(2 - n_d)} \langle \tilde{\mathcal{V}} \rangle^2. \end{aligned} \quad (60)$$

Comparing the full sum rules (34)-(36) with the projected ones (55)-(57) and also the variances, (21) and (60), one sees that the projection has two primary effects: First, it removes spectral weight associated with the UHP and the large Coulomb repulsion  $U$ . Second, it renormalizes the spectral weight from 2 (the total spectral weight) to  $\sim 2 - n_d$ , the low-energy spectral weight in the LHP. The variance (60) is essentially the variance of a *non-interacting single level*, broadened by  $\sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2$  and  $\langle \tilde{\mathcal{V}} \rangle$ ; the effects of interactions only enter through the spectral weight renormalization  $2 \rightarrow 2 - n_d$  and their effect on the coherence  $\langle \tilde{d}_\sigma^\dagger \tilde{c}_{\mathbf{k}\nu\sigma} \rangle$ . To the orders in  $V/(\epsilon_0 + U)$  shown, one can replace the transformed operators  $\tilde{d}_\sigma, \tilde{c}_{\mathbf{k}\nu\sigma}$  in the above with the original operators,  $\hat{d}_\sigma, \hat{c}_{\mathbf{k}\nu\sigma}$ , mindful, however, of the restriction that there be no doubly-occupied states.

We now use the above results to give insight into the thermoelectric performance of strongly correlated materials in the high-temperature and strongly-correlated regimes.

## V. THERMOELECTRIC PERFORMANCE AT HIGH TEMPERATURES $T \gg U$

Using the results from the previous section, we now evaluate the exact high-temperature limiting values of the thermopower  $S$  and figure of merit  $ZT$  for the Anderson impurity model. In the high temperature limit,  $n_d(T \gg U, \Gamma) \rightarrow 1$ . Using this in, (34)-(36), the high-temperature values for (15) and (16) become

$$S(T \rightarrow \infty) = \frac{\epsilon_0 - \mu + U/2}{eT} \equiv \frac{V_g}{T} \quad (61)$$

and, further using (39),

$$ZT(T \rightarrow \infty) = \frac{4V_g^2}{U^2 + 2D\Gamma/\pi}. \quad (62)$$

Here, we have introduced the *gate voltage*

$$V_g \equiv (\epsilon_0 - \mu + U/2)/e. \quad (63)$$

Taking  $D = 0$  or  $\Gamma = 0$ , (61) and (62) reproduce the analytic high-temperature limiting values found in Ref. 21 [specifically, their Eqs. (19) and (21)] using the sequential tunnelling approximation<sup>19,20</sup> for  $\Gamma \simeq 0$ . From (62), one sees that the diverging figure of merit of Mahan and

Sofo (valid for all temperatures) is realized in the non-interacting  $U \ll \epsilon_0$  and *un-broadened* limit  $D, \Gamma \ll \epsilon_0$ . The high- $T$  limiting form (61) for the Seebeck coefficient has the same form as the “Mott-Heikes” approximation  $S(T \rightarrow \infty) = [\mu(T=0) - \mu]/eT$ —well-known in studies of bulk narrow-band systems<sup>23,42</sup>—with the identification  $\mu(T=0) = \epsilon_0 + U/2$ , which is indeed the zero-temperature chemical potential at half-filling,  $n_d = 1$ .

Although these high-temperature results are valid in a temperature regime well-outside that found in typical systems, they can serve as useful checks of numerical calculations. Any conserving calculation<sup>25</sup> results in a transmission function (single-particle Green’s function) that satisfies the sum rules and hence, (61) and (62). As we have already noted, the sequential tunnelling approximation used in Ref. 21 satisfies these asymptotic results at high temperatures. Our results also give analytic expressions for the high-temperature asymptotes found numerically in Ref. 3 using numerical renormalization group<sup>43</sup>.

## VI. THERMOELECTRIC PERFORMANCE IN THE STRONGLY CORRELATED REGIME $U \gg T$ .

In the strong-correlation, narrow-level regime  $U \gg T \gtrsim \Gamma$ , (55)–(57) give

$$S(U \gg T \gtrsim \Gamma) = \frac{1}{eT} \left[ \epsilon_0 + \frac{\langle \hat{\mathcal{V}}^\dagger \rangle}{2 - n_d} - \mu + \mathcal{O}\left(\frac{V^2}{\epsilon_0 + U}\right) \right] \quad (64)$$

for the thermopower and

$$ZT = \frac{(\epsilon_0 - \mu)^2 (2 - n_d)^2 + 2(\epsilon_0 - \mu) \langle \hat{\mathcal{V}}^\dagger \rangle (2 - n_d) + \langle \hat{\mathcal{V}}^\dagger \rangle^2}{(2 - n_d)^2 \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2 + (2 - n_d) [\langle \hat{\mathcal{V}}^\dagger \hat{\mathcal{V}} \rangle - \langle \hat{\mathcal{V}}^\dagger \xi \rangle] - \langle \hat{\mathcal{V}}^\dagger \rangle^2 + \mathcal{O}\left(\frac{V}{\epsilon_0 + U}\right)}, \quad (65)$$

for the figure of merit, where we have defined ( $\xi \equiv \epsilon - \mu$ )

$$\langle \hat{\mathcal{V}}^\dagger \xi \rangle \equiv \sum_{\mathbf{k}\sigma\nu} V_{\mathbf{k}\nu\sigma}^* (\epsilon_{\mathbf{k}\nu} - \mu) \langle \hat{d}_\sigma^\dagger \hat{c}_{\mathbf{k}\nu\sigma} \rangle. \quad (66)$$

Note that we are using the bare electron operators in the above and  $\langle \hat{\mathcal{V}} \rangle$  is defined using (59). Although (64) and (65) are asymptotically exact in the limit where there is a large separation between energy scales,  $\Gamma \ll T \ll U$  (also the limit where thermoelectric performance is optimized), as noted earlier, we expect them to provide a reasonable approximation when  $\Gamma \lesssim T$ .

The thermopower (64) in the strongly-correlated regime and intermediate temperatures ( $T \gtrsim \Gamma$ ) again has the Mott-Heikes form, but with a renormalized low-energy measure  $\epsilon_0 + \langle \hat{\mathcal{V}}^\dagger \rangle / (2 - n_d)$  of the single-particle energy that does not include the Coulomb energy scale. In fact, apart from the renormalization factor  $2 - n_d$  accounting for the absence of doubly-occupied states, this

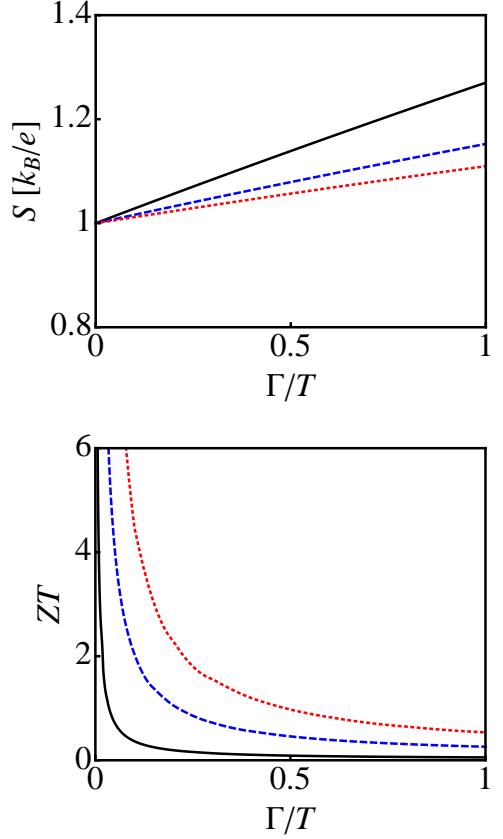


FIG. 3: Dependence of the Seebeck coefficient (top) and the electron contribution to the figure of merit (bottom) on the broadening  $\Gamma$  for different lead bandwidths:  $D/T = 100$  (black solid line),  $D/T = 20$  (blue dashed line), and  $D/T = 10$  (red dotted line). Curves are shown for  $U = 100T$  and  $\epsilon_0 - \mu = T$  although the behaviour shown here is not sensitive to the precise choice of values.

is precisely the single-particle energy of a noninteracting electron in an impurity coupled to leads by (9). Consistent with the findings of Ref. 21, by eliminating the possibility of having a doubly occupied impurity, extremely strong correlations have effectively reduced the transport problem to that of a single-level. As a result, transport is effectively that of noninteracting electrons, albeit with a reduced Hilbert space ( $n_d \neq 2$ )

As with noninteracting electrons, thermoelectric performance in this regime is limited by the broadening due to the coupling between the impurity and the lead. Specifically, when the lead-impurity coupling is larger than the detuning from the Fermi level,  $\langle \hat{\mathcal{V}} \rangle \gtrsim \epsilon_0 - \mu$ , the figure of merit (65) is generically of order unity. It is only when the coupling becomes much smaller than the detuning,  $\langle \hat{\mathcal{V}} \rangle \ll \epsilon_0 - \mu$ , that one obtains the diverging electronic contribution to the figure of merit found in Ref. 21. Outside this limit, even though the strong Coulomb repulsion has reduced the problem to an effectively single-level one, the variance in the single-particle energy (roughly,

the greater of  $\sim \sum_{\mathbf{k}\nu} |V_{\mathbf{k}\nu}|^2$  and  $\langle \hat{\mathcal{V}}^\dagger \hat{\mathcal{V}} \rangle - \langle \hat{\mathcal{V}}^\dagger \rangle^2$  prevents the system from attaining optimal thermoelectric performance.

All terms in the denominator of (65) vanish in the  $\Gamma \rightarrow 0$  or  $D \rightarrow 0$  limit, meaning that—barring a nonanalytic dependence on either of these variables—the figure of merit will exhibit the scaling shown in (2) for small  $\Gamma, D$ . To check this, in Fig. 3 we show (64) and (65) as functions of  $\Gamma$  using Hartree–Fock (HF) theory and (39) for a range of bandwidths. Details of the HF calculations are given in Appendix B. Although HF is not a quantitatively reliable theory when interactions are strong, we emphasize that we are only applying HF theory to evaluate the asymptotically exact (conserving) results (64) and (65), and we expect it to give qualitatively reliable results. We are *not* presenting the results of a (non-conserving) HF theory of transport, the results of which would be completely unreliable. While the thermopower is generically of order unity (in units of  $k_B/e$ ) for  $\epsilon_0 - \mu = T$ , the figure of merit is singular at small broadenings and is well-described by the scaling in (2). For comparison with this plot, in molecular systems,  $\Gamma \sim (0.01 \rightarrow 0.5) \text{eV}^{6,29}$  or,  $\Gamma = (0.4 \rightarrow 19) T_{\text{room}}$ , with  $T_{\text{room}} = 300K$ . For leads constructed of a good metal such as gold,  $D \sim 1 \text{eV} \sim 40 T_{\text{room}}$ .

Fig. 3 emphasizes that either  $\Gamma$  or the lead bandwidth  $D$  need to be much smaller than temperature in order to reproduce the substantial figure of merit found in Ref. 21 for  $U \rightarrow \infty$ . Reinforcing the fact that the thermoelectric transport properties in the very strongly correlated regime  $U \gg T$  is equivalent to that of a renormalized noninteracting system, we note that Fig. 3 closely resembles the  $U = 0$  Fig. 1 in Ref. 21.

## VII. DISCUSSION

Much of the interest in studying the thermoelectric performance of quantum dots and molecules originates in the likely small phonon contribution to the thermal conductance in these systems and also the fact that they can be well-approximated as single-level systems. For a single-level noninteracting impurity system, the figure of merit diverges in the limit that the extrinsic broadening arising from the coupling to the leads is much smaller than the temperature. Turning on a Coulomb repulsion, the appearance of a satellite Hubbard peak leads to an increase in the variance of the single-particle energy level and hence, a diminished figure of merit.

Somewhat counterintuitively, when the Coulomb repulsion is *very* strong, however, the new satellite Hubbard level becomes unimportant, as it is too high an energy scale to be relevant for transport. We have used asymptotically exact sum rule expressions in this limit to show that the thermoelectric transport coefficients assume essentially the same form as for a noninteracting system, slightly renormalized to account for the absence of doubly occupied states. The thermopower (64) in par-

ticular assumes the well-known Mott–Heikes form, but with an effective noninteracting chemical potential. As with a noninteracting system, the transport properties in this strong interaction regime are limited only by the broadening due to the coupling between the leads and the impurity.

This bodes well for using small impurity-type systems such as quantum dots and molecules to achieve substantial thermoelectric efficiencies. Even though the bare (un-screened) Coulomb energy scale is very large ( $U \gg T$ ) in these systems, it is precisely its large size that means it does not lead to a diminishment of transport, as would happen if it was of moderate strength, say  $U \sim T$ . At the same time, the highly singular dependence of the figure of merit on the broadening and also the lead bandwidth [c.f. (2)] means that it will be essential to have a small broadening arising from lead-impurity coupling or to tailor the bandwidth properties of the leads in order to realize the potential of impurity systems.

## Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council of Canada Discovery Grant and the Canada Research Chair Program.

## Appendix A: Spectral properties of the single-particle Green's function

In this Appendix, we give an example of the derivation of the general result (25) for  $n = 1$ :

$$\begin{aligned}
M_1^T &\equiv \int_{-\infty}^{\infty} dEEA(E) \\
&= \sum_{\sigma} \sum_{a,b} (P_a + P_b) \langle a | \hat{d}_{\sigma} | b \rangle \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle (E_b - E_a) \\
&= \sum_{\sigma} \sum_{a,b} P_a \left[ \langle a | \hat{d}_{\sigma} \hat{H} | b \rangle \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle - \langle a | \hat{H} \hat{d}_{\sigma} | b \rangle \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle \right] \\
&\quad + \sum_{\sigma} \sum_{a,b} P_b \left[ \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle \langle a | \hat{d}_{\sigma} \hat{H} | b \rangle - \langle b | \hat{d}_{\sigma}^{\dagger} | a \rangle \langle a | \hat{H} \hat{d}_{\sigma} | b \rangle \right] \\
&= \sum_{\sigma} \sum_a P_a \left[ \langle a | \hat{d}_{\sigma} \hat{H} \hat{d}_{\sigma}^{\dagger} | a \rangle - \langle a | \hat{H} \hat{d}_{\sigma} \hat{d}_{\sigma}^{\dagger} | a \rangle \right] \\
&\quad + \sum_{\sigma} \sum_b P_b \left[ \langle b | \hat{d}_{\sigma}^{\dagger} \hat{d}_{\sigma} \hat{H} | b \rangle - \langle b | \hat{d}_{\sigma}^{\dagger} \hat{H} \hat{d}_{\sigma} | b \rangle \right] \\
&= \sum_{\sigma} \left[ \langle \hat{d}_{\sigma} \hat{H} \hat{d}_{\sigma}^{\dagger} \rangle - \langle \hat{H} \hat{d}_{\sigma} \hat{d}_{\sigma}^{\dagger} \rangle + \langle \hat{d}_{\sigma}^{\dagger} \hat{d}_{\sigma} \hat{H} \rangle - \langle \hat{d}_{\sigma}^{\dagger} \hat{H} \hat{d}_{\sigma} \rangle \right] \\
&= \sum_{\sigma} \langle \{ \hat{d}_{\sigma}^{\dagger}, [\hat{d}_{\sigma}, \hat{H}] \} \rangle. \tag{A1}
\end{aligned}$$

## Appendix B: Hartree–Fock theory for the AIM model

In this Appendix, we describe how to calculate the impurity occupation  $n_d$ , lead-impurity coherence  $\langle \hat{\mathcal{V}} \rangle =$

$\langle \hat{\mathcal{V}}^\dagger \rangle$ , and  $\langle \tilde{\mathcal{V}}^\dagger \xi \rangle$ , defined in (66) using Hartree–Fock (HF) theory. Within this approximation,  $\langle \hat{\mathcal{V}}^\dagger \hat{\mathcal{V}} \rangle = \langle \hat{\mathcal{V}} \rangle^2$  and hence, these three quantities are sufficient to evaluate the figure of merit (65).

Momentarily generalizing the AIM hamiltonian (32) to allow for different chemical potentials and couplings for each spin,  $\mu \rightarrow \mu_\sigma$  and  $V_{\mathbf{k}\nu} \rightarrow V_{\mathbf{k}\nu\sigma}$ , one has

$$n_\sigma \equiv \langle \hat{n}_\sigma \rangle = - \left( \frac{\partial \Omega}{\partial \mu_\sigma} \right)_T \quad (\text{B1})$$

and

$$\langle \hat{d}_\sigma \hat{c}_{\mathbf{k}\nu\sigma}^\dagger \rangle = - \langle \hat{c}_{\mathbf{k}\nu\sigma}^\dagger \hat{d}_\sigma \rangle = - \left( \frac{\partial \Omega}{\partial V_{\mathbf{k}\nu\sigma}} \right)_T, \quad (\text{B2})$$

where  $\Omega$  is the free energy. At the end of our calculation, we set  $\mu_\uparrow = \mu_\downarrow = \mu$  and likewise with  $V_{\mathbf{k}\nu\sigma}$ . (B1) is a standard thermodynamic identity, while (B2) follows from an application of the Hellmann–Feynman formula  $(\partial \Omega / \partial \lambda)_{T,N} = (\partial \hat{H} / \partial \lambda)$  with  $\lambda = V_{\mathbf{k}\nu\sigma}$  to (32).

Since it is an effective single-particle theory, the HF free energy is given by the trace of the logarithm of the HF Green's function  $G_\sigma(i\omega_n)$  for spin  $\sigma$  impurity electrons:  $\Omega = -\beta^{-1} \sum_{n\sigma} e^{i\omega_n 0^+} \ln[-G_\sigma^{-1}(i\omega_n)]$ , yielding

$$\Omega = -\frac{1}{\beta} \sum_{n\sigma} e^{i\omega_n 0^+} \ln \left[ -i\omega_n + \epsilon_0 - \mu_\sigma + Un_\sigma + \sum_{\mathbf{k}\nu} \frac{|V_{\mathbf{k}\nu\sigma}|^2}{i\omega_n - \xi_{\mathbf{k}\nu\sigma}} \right]. \quad (\text{B3})$$

Here  $\omega_n = (2n + 1)/\beta$  are Fermi Matsubara frequencies, with integer  $n$ , and  $\xi_{\mathbf{k}\nu\sigma} \equiv \epsilon_{\mathbf{k}\nu} - \mu_\sigma$ . The  $\exp(i\omega_n 0^+)$  factor with positive infinitesimal  $0^+$  ensures convergence. Making the same assumptions (constant lead DOS, momentum-independent coupling  $V_{\mathbf{k}\nu\sigma}$ ) that we used to arrive at (38), (B3) reduces to

$$\Omega = -\frac{1}{\beta} \sum_{n\sigma} e^{i\omega_n 0^+} \ln \left[ -i\omega_n + \epsilon_0 - \mu_\sigma + Un_\sigma - \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right) \right]. \quad (\text{B4})$$

We will not make the usual approximation in the AIM literature of taking  $D \rightarrow \infty$ , in which case  $\ln[(i\omega_n - D)/(i\omega_n + D)]$  reduces to  $\pi \text{sgn}(\omega_n)$  (see e.g., Sec. 5.2 in Ref. 35).

Applying (B1) and (B2) to (B3) [for the coherence (B2) we need to take the derivative of the free energy with respect to  $V_{\mathbf{k}\nu\sigma}$  before assuming the coupling to be

constant], one arrives at the results

$$n_\sigma = \frac{1}{\beta} \sum_n \frac{e^{i\omega_n 0^+}}{i\omega_n - \epsilon_0 + \mu_\sigma - Un_\sigma + \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right)} \quad (\text{B5})$$

and

$$\langle \hat{d}_\sigma \hat{c}_{\mathbf{k}\nu\sigma}^\dagger \rangle = -\frac{1}{\beta} \sum_n \frac{e^{i\omega_n 0^+}}{(i\omega_n - \xi_{\mathbf{k}\sigma})} \times \frac{V_{\mathbf{k}\nu\sigma}^*}{\left[ i\omega_n - \epsilon_0 + \mu_\sigma - Un_\sigma + \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right) \right]} \quad (\text{B6})$$

Using this last result to evaluate  $\langle \hat{\mathcal{V}} \rangle$  and (66) and replacing the momentum integration with one over energy [assuming again a constant lead DOS as in (37)] gives

$$\langle \hat{\mathcal{V}} \rangle = \frac{1}{\beta} \sum_n \frac{e^{i\omega_n 0^+} \left( \frac{\Gamma}{2\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right)}{i\omega_n - \epsilon_0 + \mu_\sigma - Un_\sigma + \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right)} \quad (\text{B7})$$

and

$$\langle \tilde{\mathcal{V}}^\dagger \xi \rangle = \frac{1}{\beta} \sum_n \frac{e^{i\omega_n 0^+} \left( \frac{\Gamma}{2\pi} \right) \left[ 2D + i\omega_n \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right) \right]}{i\omega_n - \epsilon_0 + \mu_\sigma - Un_\sigma + \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right)}, \quad (\text{B8})$$

To obtain the results shown in Fig. 3, we solve (B5) self-consistently for  $n_\sigma$ , evaluating the Matsubara frequency sum numerically by summing between  $n = -10^4 \rightarrow 10^4$ . To avoid dealing with the convergence factor, we use

$$n_\sigma = \frac{1}{2} + \frac{1}{\beta} \sum_n \frac{1}{i\omega_n - \epsilon_0 + \mu_\sigma - Un_\sigma + \left( \frac{\Gamma}{4\pi} \right) \ln \left( \frac{i\omega_n - D}{i\omega_n + D} \right)}. \quad (\text{B9})$$

This can be derived from (B5) by re-writing the frequency sum as an integral over the complex plane and manipulating the contour of integration; see e.g. Appendix A in Ref. 44. In contrast to (B5), for which the convergence factor is needed to deal with a summand that vanishes as  $(i\omega_n)^{-1}$  at large frequencies, the summands in both (B7) and (B8) vanish as  $(i\omega_n)^{-2}$ , and we can simply drop the convergence factor. The solution of (B9) is used in these and we again sum over the Matsubara frequencies, from  $n = -10^4 \rightarrow 10^4$ , to obtain  $\langle \hat{\mathcal{V}} \rangle$  and  $\langle \tilde{\mathcal{V}}^\dagger \xi \rangle$ .

<sup>1</sup> G. D. Mahan and J. O. Sofo, Proc. Nat. Acad. Sci. **93**, 7436 (1996).

<sup>2</sup> R. Scheibner, H. Buhmann, D. Reuter, M. N. Kiselev, and L. W. Molenkamp, Phys. Rev. Lett. **95**, 176602 (2005).

<sup>3</sup> T. A. Costi and V. Zlatić, Phys. Rev. B **81**, 235127 (2010).

<sup>4</sup> F. Haupt, M. Leijnse, H. L. Calvo, L. Classen, J. Splettstoesser, and M. R. Wegewijs, Phys. Status Solidi B **250**, 2315 (2029).

<sup>5</sup> B. Sothmann, R. Sánchez, and A. N. Jordan,

Nanotechnology **26**, 032001 (2015).

<sup>6</sup> M. Paulsson and S. Datta, Phys. Rev. B **67**, 241403(R) (2003).

<sup>7</sup> P. Reddy, S.-Y. Jang, R. A. Segalman, and A. Majumdar, Science **315**, 1568 (2007).

<sup>8</sup> J. A. Malen, S. K. Yee, A. Majumdar, and R. A. Segalman, Chem. Phys. Lett. **491**, 109 (2010).

<sup>9</sup> S. Andergassen, T. A. Costi, and V. Zlatić, Phys. Rev. B **84**, 241107(R) (2011).

<sup>10</sup> Y. Dubi and M. Di Ventra,

Rev. Mod. Phys. **83**, 131 (2011).

<sup>11</sup> S. K. Yee, J. A. Malen, A. Majumdar, and R. A. Segalman, Nano Lett. **11**, 4089 (2011).

<sup>12</sup> W. Lee, K. Kim, W. Jeong, L. A. Zotti, F. Pauly, J. C. Cuevas, and P. Reddy, Nature **498**, 209 (2014).

<sup>13</sup> Y. Kim, W. Jeong, K. Kim, W. Lee, and P. Reddy, Nature Nanotech. **9**, 881 (2014).

<sup>14</sup> R. Stadler and T. Markussen, J. Chem. Phys. **135**, 154109 (2011).

<sup>15</sup> M. Bürkle, T. J. Hellmuth, F. Pauly, and Y. Asai, Phys. Rev. B **91**, 165419 (2015).

<sup>16</sup> T. A. Costi, A. C. Hewson, and V. Zlatić, J. Phys. Condens. Matter **6**, 2519 (1994).

<sup>17</sup> J. Liu, Q.-f. Sun, and X. C. Xie, Phys. Rev. B **81**, 245323 (2010).

<sup>18</sup> J. Ren, J.-X. Zhu, J. E. Gubernatis, C. Wang, and B. Li, Phys. Rev. B **85**, 155443 (2012).

<sup>19</sup> C. W. J. Beenakker and A. A. M. Staring, Phys. Rev. B **46**, 9667 (1992).

<sup>20</sup> J. Koch, F. von Oppen, Y. Oreg, and E. Sela, Phys. Rev. B **70**, 195107 (2004).

<sup>21</sup> P. Murphy, S. Mukerjee, and J. Moore, Phys. Rev. B **78**, 161406(R), (2008).

<sup>22</sup> G. Beni, Phys. Rev. B **10**, 2186 (1974).

<sup>23</sup> P. M. Chaikin and G. Beni, Phys. Rev. B **13**, 647 (1976).

<sup>24</sup> S. Mukerjee and J. E. Moore, App. Phys. Lett. **90**, 112107 (2007).

<sup>25</sup> G. Baym and L. P. Kadanoff, Phys. Rev. **124**, 287 (1961).

<sup>26</sup> G. Benenti, G. Casati, T. Prosen, and K. Saito, arXiv:1311.4430.

<sup>27</sup> Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).

<sup>28</sup> T.-S. Kim and S. Hershfield, Phys. Rev. Lett. **88**, 136601 (2002).

<sup>29</sup> S. Y. Quek, L. Venkataraman, H. J. Choi, S. G. Louie, M. S. Hybertsen, and J. B. Neaton, Nano Lett. **7**, 3477 (2007).

<sup>30</sup> S. R. White, Phys. Rev. B **44**, 4670 (1991).

<sup>31</sup> A. B. Harris and R. V. Lange, Phys. Rev. **157**, 295 (1967).

<sup>32</sup> This bound actually furnishes a much stronger bound for the thermal conductance  $\kappa_0$  at zero chemical potential (voltage) difference,  $\mu_L = \mu_R$ , since  $\kappa_0 = (\Gamma/8\hbar T^2)M_2$ .

<sup>33</sup> E. Taylor and D. Segal, Phys. Rev. Lett. **114**, 220401 (2015).

<sup>34</sup> By “negligible” here, we mean  $E^2 A(E) \ll T^2$  for  $\Gamma \ll E \ll U$ , since this quantity enters the  $n = 2$  sum rules. Since the high-frequency tail  $\lim_{E \gg U} A(E) \propto E^{-2}$  recovers the noninteracting Lorentzian asymptote, one also requires the lead bandwidth  $D$  to be  $\lesssim U$  in order to satisfy this constraint.

<sup>35</sup> A. C. Hewson, *The Kondo Problem To Heavy Fermions*, Cambridge Studies in Magnetism, (Cambridge University Press, Cambridge, England, 1997).

<sup>36</sup> For examples of sum-rule calculations involving non-standard weight functions using the OPE, see: E. C. Poggio, H. R. Quinn, and S. Weinberg, Phys. Rev. D **13**, 1958 (1976) and E. Braaten, D. Kang, and L. Platter, Phys. Rev. Lett. **104**, 223004 (2010).

<sup>37</sup> H. Eskes, A. M. Oleś, M. B. J. Meinders, and W. Stephan, Phys. Rev. B **50**, 17980 (1994).

<sup>38</sup> M. Randeria, R. Sensarma, N. Trivedi, and F.-C. Zhang, Phys. Rev. Lett. **95**, 137001 (2005).

<sup>39</sup> J. R. Schrieffer and P. A. Wolff, Phys. Rev. **149**, 491 (1966).

<sup>40</sup> A. H. MacDonald, S. M. Girvin, and D. Yoshioka, Phys. Rev. B **37**, 9753 (1988).

<sup>41</sup> This result only differs from the Schrieffer–Wolff transformation in the form of the denominator which for us, does not include the kinetic energy of the bath electrons. This difference—unimportant for the final result—stems from our having included the corresponding term in the hamiltonian with the other hopping terms, while Schrieffer and Wolff keep them in  $\hat{H}_0$ .

<sup>42</sup> M. R. Peterson and B. S. Shastry, Phys. Rev. B **82**, 195105 (2010).

<sup>43</sup> R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

<sup>44</sup> Y. Ohashi and A. Griffin, Phys. Rev. A **67**, 033603 (2003).